ZOOLOGICAL SCIENCE 6: 269–281 (1989)

Hemoglobins from the Two Closely Related Clams *Barbatia lima* and *Barbatia virescenes*. Comparison of Their Subunit Structures and N-terminal Sequence of the Unusual Two-Domain Chain

Tomohiko Suzuki[1], Makoto Shiba, Takahiro Furukohri and Michiyori Kobayashi[2]

Department of Biology, Faculty of Science, Kochi University, Kochi 780, and [2]Department of Biology, Faculty of Science, Niigata University, Niigata 880, Japan

ABSTRACT—The intracellular hemoglobins were isolated from the two closely related arcid clams *Barbatia lima* and *Barbatia virescens*, respectively, and their subunit structures were compared in detail.

B. virescens hemoglobin on Ultrogel AcA44 column had a single peak corresponding to the molecular weight (*Mr*) of 30,000. The hemoglobin was separated by high-performance liquid chromatography (HPLC) into two chains (I and II) with *Mr* 16,000, suggesting the heterodimeric structure. N-Terminal 20 residues of chain I showed homology (35% identity) with those of *Anadara* homodimeric hemoglobins.

On the other hand, *B. lima* hemoglobin on Ultrogel AcA44 column had two peaks, a polymeric hemoglobin eluted in the void volume of the column and a *Mr* 60,000 hemoglobin. The polymeric hemoglobin on HPLC eluted as a single peak, and the *Mr* was estimated to be 32,000 by sodium dodecyl sulfate-polyacrylamide gel electrophoresis in the presence of a reducing agent. Therefore, it was suggested that the constituent chain of polymeric hemoglobin has an unusual two-domain structure with two hemes per chain. The *Mr* 60,000 hemoglobin was separated by HPLC into two major chains (A and B) with *Mr* 16,000 suggesting the tetrameric structure (A_2 B_2). These structural features of *B. lima* hemoglobins were similar to those of *B. reeveana* [1].

N-Terminal sequences of 36 residues of each domain of *B. lima* two-domain chain were determined from three CNBr fragments. N-Terminus of the chain was blocked by an acetyl group, and the second domain appeared to start at methionine residue. The sequence homology between domains were very high; 27 out of 36 residues (75%) were identical. We propose from these results that the two-domain chain resulted from gene duplication and the following loss of a stop codon.

INTRODUCTION

Several taxodont(arcid) and heterodont clams have hemoglobins in circulating erythrocytes [2]. Usually, their quaternary structure is dimeric or tetrameric, but the subunit assembly has been shown to be quite different from that of vertebrate tetrameric hemoglobin [3]. Like tetrameric hemoglobin, the dimeric hemoglobin shows a cooperativity [2]. It is also noted that the clam hemoglobins have an N-terminal extension of amino acid sequence which forms an additional helix (namely pre-A helix) [3].

Recently, Grinich and Terwilliger [1, 4] found that the arcid clam *Barbatia reeveana* has not only a tetrameric hemoglobin but also a polymeric hemoglobin in erythrocytes. Furthermore they showed that the polymeric hemoglobin appeared to be composed of unusual two-domain polypeptide chains with two hemes. Riggs *et al.* [5] reported the partial amino acid sequence (the first 129 residues of domain 1 and the middle 108

Accepted June 25, 1988
Received May 23, 1988
[1] To whom all correspondence should be addressed.

residues of domain 2) of *B. reeveana* two-domain hemoglobin by cDNA sequencing.

In this paper, we report the subunit structures of hemoglobins from *Barbatia lima* and *Barbatia virescens*. It will be shown that these closely related clams have quite different hemoglobins. We also report the N-terminal sequences of 36 residues of each domain of *B. lima* two-domain hemoglobin.

MATERIALS AND METHODS

B. lima and *B. virescens* were collected from rocks near Usa, Kochi, Japan. *B. lima* is a open-sea type species, while *B. virescens* is a bay-type species. They have a similar shell but are distinguished by the hinge line and the interior color of shell.

The red cells were washed three times with 3% (w/v) NaCl and lysed with ice-cold 1 mM phosphate buffer (pH 7.2). The hemolysate was centrifuged at 15,000 rpm for 15 min, and the supernatant was immediately applied to the column of Ultrogel AcA 44 equilibrated with 0.1 M phosphate buffer (pH 7.2).

B. virescens hemoglobin passed through a gel filtration column was applied to a DEAE-cellulose column equilibrated with 20 mM Tris-HCl (pH 8.0) and eluted with a linear gradient of 0 to 40 mM NaCl in the same buffer.

B. lima polymeric hemoglobin, which is eluted in the void volume of the gel filtration column (Fig. 1B), was applied to a DEAE-cellulose column equilibrated with 20 mM Tris-HCl (pH 8.0) containing 0.1 mM dithiothreitol(DTT) and eluted with a linear gradient of 0 to 0.1 M NaCl in the same buffer. *B. lima Mr* 60,000 hemoglobin, which corresponds to the latter peak in Fig. 1B, was also applied to a DEAE-cellulose column equilibrated with 20 mM Tris-HCl (pH 9.4) and eluted with a linear gradient of 20 mM Tris-HCl (pH 9.4) to 20 mM Tris-HCl (pH 8.0).

The constituent polypeptide chains of hemoglobin were separated by high performance liquid chromatography (HPLC, Hitachi 655) on a C_{18}-μBondapak column (Waters) using a solvent system as described by Shelton *et al.* [6], or on a Lichrosorb RP2 column using a solvent system by Ashauer *et al.* [7].

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was carried out in 15% acrylamide containing 0.1% SDS and 0.375 M Tris-HCl (pH 8.9). Samples containing 2.5% SDS and 5% 2-mercaptoethanol were incubated at 100°C for 5 min before electrophoresis.

CNBr cleavage of *B. lima* two-domain hemoglobin (370 nmoles) was performed in 70% formic acid with a 200-fold excess of CNBr over methionine at 37°C for 20 hr, and the CNBr peptides were fractionated by a Sephadex G-50 column (1.6×100 cm) equilibrated with 8.7% (v/v) acetic acid / 2.5% formic acid. The peptides were purified further by HPLC on a Cosmosil $5C_{18}$-300 column (Nakarai Chemicals Ltd.) with a linear gradient of 2 to 80% acetonitrile in 0.1% trifluoroacetic acid (TFA). Two CNBr fragments, CN1 and CN3, were further digested with trypsin (Worthington) in 0.1 M ammonium bicarbonate at 37°C for 2hr at an enzyme-to-substrate ratio of 1/50 (w/w). The tryptic peptides were purified by HPLC using the same conditions as described above. The subpeptide CN1T1 was digested with lysyl endopeptidase (Wako) in 10 mM Tris-HCl (pH 8.9) at 37°C for 2 hr at an enzyme-to-substrate ratio of 1/200, and the resultant peptides (CN1T1L1 and CN1T1L2) were separated by HPLC. To remove the acylamino-acid, CN1T1L1 was digested further with acylamino-acid-releasing enzyme (0.025 units) (Takara) in 5 mM phosphate buffer (pH 7.2) containing 2-mercaptoethanol at 37°C for 2 hr.

B. lima two-domain hemoglobin (50 nmoles) was also digested with lysyl endopeptidase at 37°C for 2.2 hr, and the peptides were purified by HPLC.

Proteins and peptides were hydrolyzed with TFA/HCl (1/2, v/v) containing 0.02% phenol at 170°C for 30 min in evacuated sealed tubes. Amino acid analysis was performed in Hitachi 835-50 amino acid analyzer.

The amino acid sequence was determined by the manual Edman method with modification [8, 9]. Phenylthiohydantoin amino acid derivatives were identified by HPLC on a Cosmosil 5PTH column (Nakarai Chemicals Ltd.) with isocratic elution.

The primary peptides were numbered from the

amino-terminus, and the subpeptides are numbered in order within the parent pepide. In this paper, the prefix CN indicates a cyanogen bromide peptide; T, a tryptic peptide; L, a lysyl endopeptidase peptide; and A, a acylamino-acid-releasing enzyme peptide.

RESULTS

Gel filtration profiles of *B. virescens* and *B. lima* red cell hemolysates are shown in Figs. 1A and 1B, respectively. *B. virescens* hemoglobin had a single peak corresponding to the molecular weight (*Mr*) of 30,000, while *B. lima* hemoglobin had two peaks; one is a polymeric hemoglobin eluted in the void volume of the column and the other is a *Mr* 60,000 hemoglobin. Although we did not estimate the *Mr* of polymeric hemoglobin of *B. lima,* Grinich and Terwilliger [1] reported the value of 430,000 to the polymeric hemoglobin of *B. reeveana.*

DEAE-cellulose chromatography of *B. virescens Mr* 30,000 hemoglobin and *B. lima* polymeric and *Mr* 60,000 hemoglobins gave one major peak containing heme, respectively (Figs. 2A-C).

Figure 3 shows the SDS-PAGE patterns of reduced hemoglobins from *B. virescens* and *B. lima.* The hemolysate and purified hemoglobin (lanes 1 and 2) of *B. virescens* gave a single band (*Mr* 16,000). On the other hand, the hemolysate (lane 3) of *B. lima* gave two bands (*Mr*s 32,000 and 16,000), which corresponded to polymeric (lane 4) and *Mr* 60,000 (lane 5) hemoglobins, respectively.

B. virescens hemoglobin was separated by HPLC into two constituent polypeptide chains (I and II) in equimolar proportions (Fig. 4A). Amino acid compositions of isolated chains are shown in Table I. N-Terminal sequence (20 residues) of chain I was determined to be Pro-Ser-Val-Ala-Ala-Ala-Val-Ser-Ala-Val-Thr-Asn-Lys-Asp-Val-Ala-Gln-Glu-Ile-Trp- by manual Edman degradation, but that of chain II was not detected, suggesting that N-terminus is blocked. The *Mr*s of chains I and II were estimated to be 16,000 by SDS-PAGE. Based on gel filtration, DEAE-cellulose chromatography and SDS-PAGE, we concluded that *B. virescens* hemoglobin has a heterodimeric structure.

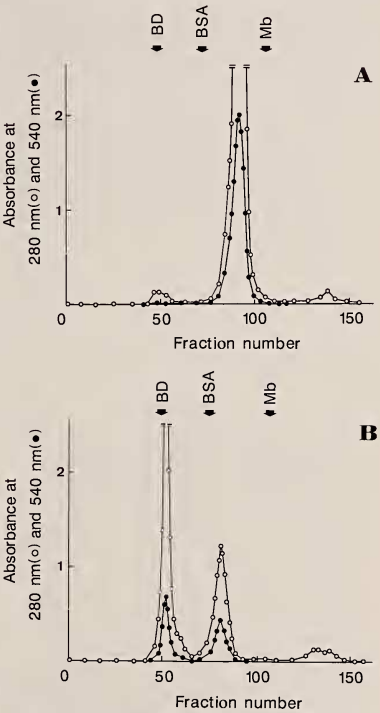


FIG. 1. (A) Gel filtration on Ultrogel AcA 44 of *B. virescens* red cell hemolysate. The column (3× 100 cm) was equilibrated with 0.1 M phosphate buffer (pH 7.2). Fraction size, 5 ml/tube. BD, blue dextran; BSA, bovine serum albumin; Mb, sperm whale myoglobin. (B) Gel filtration on Ultrogel AcA 44 of *Barbatia lima* red cell hemolysate. Conditions are as for (A).

B. lima Mr 60,000 hemoglobin was separated by HPLC into two major chains (A and B) and a minor chain (A') as shown in Fig. 4B. Amino acid composition of chain A' was very similar to that of chain A (Table 1), suggesting that chain A' is a hetero-type of chain A. No N-terminal residue of all the chains was detected by manual Edman method. The *Mr*s of chains were determined to be

16,000 by SDS-PAGE. Based on gel filtration, DEAE-cellulose chromatography and SDS-PAGE, we concluded that *B. lima Mr* 60,000 hemoglobin has a heterotetrameric structure (A_2B_2).

B. lima polymeric hemoglobin was purified by HPLC on a Lichrosorb RP2 column (Fig. 4C) or a C_{18}-μBondapak column (data not shown). In eash case, only one polypeptide chain (P) was eluted. No N-terminal residue was detected by manual Edman method. The *Mr* of chain P was detemined

FIG. 3. SDS-PAGE patterns of reduced *Barbatia* hemoglobins. Lane 1, red cell hemolysate of *B. virescens;* 2, purified hemoglobin (Fig. 1A) of *B. virescens;* 3, red cell hemolysate of *B. lima;* 4, polymeric hemoglobin (Fig. 1B) of *B. lima;* 5, *Mr* 60,000 hemoglobin (Fig. 1B) of *B. lima.*

to be 32,000 by SDS-PAGE in the presence of reducing agent, which is two times as large as that of a typical globin subunit. Therefore we concluded that *B. lima* polymeric hemoglobin consists of one major chain and that the chain has an unusual two-domain structure with two hemes, like *B. reeveana* polymeric hemoglobin [1].

B. lima two-domain chain was cleaved with CNBr, and the CNBr peptides were fractionated by Sephadex G-50 chromatography (Fig. 5). Three peptides, CN1 (25 residues), CN2 (11 re-

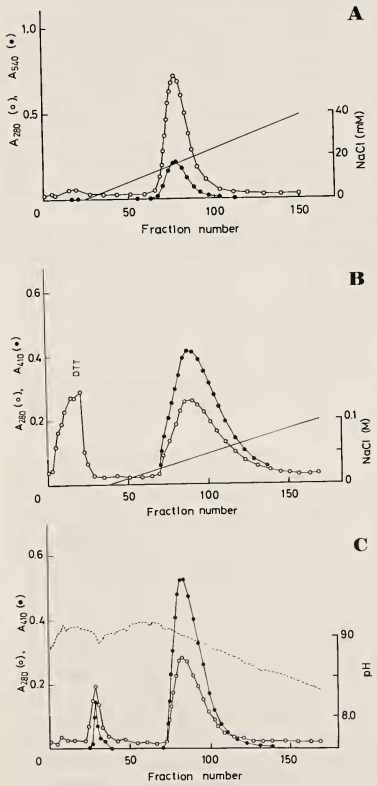
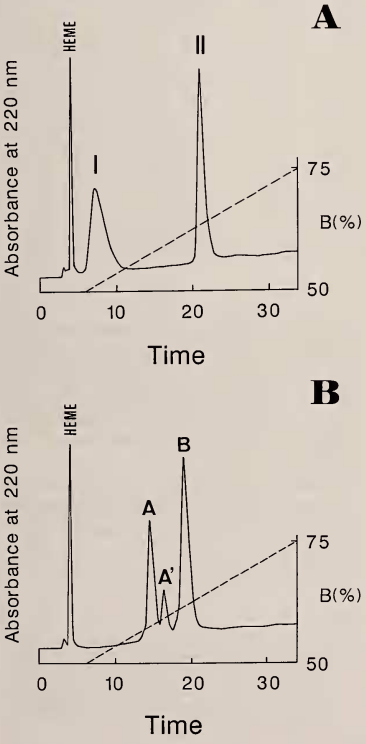


FIG. 2. (A) DEAE-cellulose chromatography of *B. virescens Mr* 30,000 hemoglobin. The column (1.3×10 cm) was equilibrated with 20 mM Tris-HCl (pH 8.0) and eluted with a linear gradient of 0 to 40 mM NaCl in the same buffer. Fraction size, 5 ml/tube. (B) DEAE-cellulose chromatography of *B. lima* polymeric hemoglobin. The column (1.3×10 cm) was equilibrated with 20 mM Tris-HCl (pH 8.0) containing 0.1 mM DTT and eluted with a linear gradient of 0 to 0.1 M NaCl in the same buffer. The sample was reduced with 2 mM DTT before chromatography. Fraction size, 5 ml/tube. (C) DEAE-cellulose chromatography of *B. lima Mr* 60,000 hemoglobin. The column (1.3×10 cm) was equilibrated with 20 mM Tris-HCl (pH 9.4) and eluted with a linear gradient of 20 mM Tris-HCl (pH 9.4) to 20 mM Tris-HCl (pH 8.0). Fraction size, 5 ml/tube.

sidues) and CN3 (36 residues), were purified further by HPLC (Fig. 6). CN1 with blocked N-terminal residue was placed at the N-terminus of the whole protein. CN3 had strong sequence homology with CN1, therefore, we placed CN3 at the N-terminus of the second domain. With the aid of sequence homology, CN2 was placed behind CN1.

An N-terminal acetyl group was determined as dansylacetylhydradine [8], using the peptide CN1.

CN1 and CN3 were digested with trypsin, and the tryptic peptides were purified by HPLC (Fig. 7). CN1 produced three tryptic peptides, CN1T1 with blocked N-terminus, CN1T2 and CN1T3 with



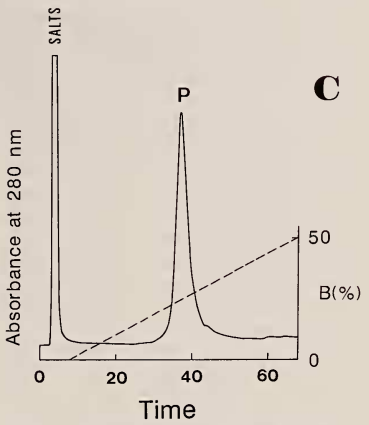


FIG. 4. (A) Separation of globin chins of *B. virescens* hemoglobin by HPLC. The column (C_{18}-μBondapak) was equilibrated with 50% solvent A/50% solvent B and eluted with a linear gradient of 50 to 75% solvent B over 25 min. Solvent A in volume %, 80:5:15:0.1:0.05 of 0.15 M $NaClO_4$: methanol: acetonitrile: 85% H_3PO_4: nonylamine. Solvent B, 20:5:75:0.1:0.05 of the same components. Flow rate, 1 ml/min. (B) Separation of globin chins of *B. lima Mr* 60,000 hemoglobin by HPLC. Conditions are as for (A). (C) Separation of globin chain of *B. lima* polymeric hemoglobin by HPLC. Heme was removed before HPLC, and the apopotein was reduced with 10 mM DTT. The column (Lichrosorb RP2) was equilibrated with solvent A and eluted with a linear gradient of 0 to 100% solvent B over 60 min. Solvent A, 50 mM ammonium acetate containing 5% HCOOH and 15% acetonitrile. Solvent B, 50 mM ammonium acetate containing 5% HCOOH and 70% acetonitrile. Flow rate, 0.7 ml/min.

homoserine residue. Therefore we aligned them in the order CN1T1-CN1T2-CN1T3. CN1T1 was digested further with lysyl endopeptidase, and the N-terminal peptide CN1T1L1 was isolated. Since the N-terminus of CN1T1L1 is blocked by an acetyl group, we digested this peptide with acylamino-acid-releasing enzyme to remove the acylamino acid (acetylserine) and then sequenced the resultant peptide without trouble. CN3 produced

TABLE 1. Amino acid compositions of constituent chains of *Barbatia* hemoglobins

	B. virescens Dimeric		*B. lima* Tetrameric			*B. lima* Polymeric
	I	II	A	A'	B	P
Asp	16.6	18.6	16.6	16.2	18.9	43.4
Thr	6.3	3.7	8.8	8.3	5.6	9.0
Ser	9.5	9.0	9.8	10.9	8.7	6.2
Glu	15.6	11.8	13.6	14.0	8.5	31.1
Pro	5.9	3.8	3.0	3.4	4.7	8.7
Gly	6.9	9.7	12.5	12.8	7.4	19.8
Ala	12.7	15.5	11.9	13.5	22.6	23.1
Cys	0.0	0.3	0.9	1.1	1.3	1.8
Val	10.6	8.2	8.5	7.9	9.3	24.8
Met	5.6	2.8	6.7	6.0	4.7	6.0
Ile	5.2	7.3	5.9	5.6	5.2	14.2
Leu	14.6	18.3	16.5	15.6	15.8	33.2
Tyr	1.9	5.6	1.2	1.8	4.3	3.1
Phe	8.7	5.8	7.4	7.2	6.3	10.4
Lys	11.5	15.0	11.1	10.1	12.1	29.4
His	2.0	3.1	2.5	2.9	3.1	6.9
Arg	10.5	5.9	8.0	7.8	6.8	18.7
Trp.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Total	144	144	145	145	145	290
SΔQ	43.3	83.4	50.9	42.0	131.3	98.9

N.D. not determined.

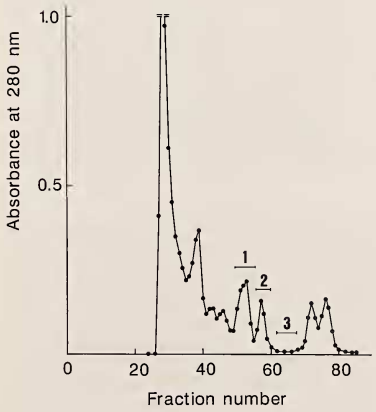
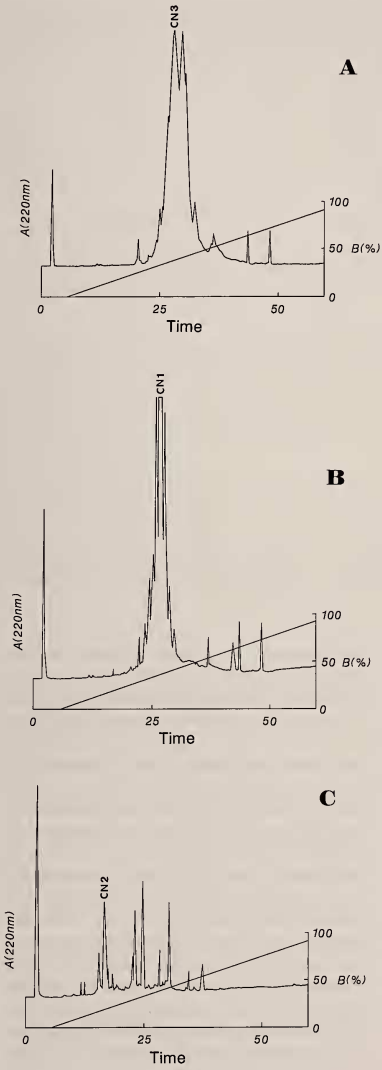


FIG. 5. Gel filtration on Sephadex G50 of CNBr fragments of *B. lima* two-domain hemoglobin. The column (1.6×100 cm) was equilibrated with 8.7% acetic acid/2.5% formic acid. Three fractions (1, 2 and 3) were pooled, respectively. Fraction size, 2 ml/tube.

five tryptic peptides. Since we determined the first 14 residues of CN3 by the direct sequencing, CN1T1 and CN1T2 were easily placed in this region. CN3T5 with homoserine residue was placed at the C-terminus, and the remaining two peptides (CN3T3 and CN3T4) were aligned by the homology with CN1.

B. lima two-domain chain was also digested with lysyl endopeptidase, and the N-terminal peptides L1 and L2 of second domain were purified by HPLC (Fig. 8).

The amino acid compositions of peptides are



shown in Table 2.

The amino acid sequence was mainly determined from the tryptic peptides of CNBr fragments. The procedures used for sequence determination are summarized in Fig. 9. N-Terminal Ser of domain 1 was determined from amino acid composition of the peptide CN1T1L1. Val-Met at positions 24–25 (domain 1), Leu-Met at 35–36 (domain 1), Lys at 30 (domain 2) and Leu-Met at 35–36 (domain 2) were determined from amino acid compositions of peptides and specificity of CNBr cleavage or trypsin. Heterogeneity was found at positions 10 (Thr and Ala) and 29 (Arg and Lys) of domain 2. In each case, the former amino acid in parentheses was obtained at high yield.

DISCUSSION

One of the recent topics on molluscan hemoglobins was the finding of polymeric hemoglobin (*Mr* 430,000) in the circulating erythrocytes of *B. reeveana* [1]. The hemoglobin was not only the largest intracellular hemoglobin so far known, but also consisted of unusual two-domain polypeptide chains (*Mr* 32,000–34,000). In order to elucidate the physiological role and evolutionary origin of the unusual two-domain chain, we examined the subunit structures of the hemoglobins from *B. lima* and *B. virescens* in detail.

B. virescens contained only a heterodimeric hemoglobin consisting of myoglobin-like subunits. This seems rather unique, since many arcids such as *Anadara* [10, 11] and *Scapharca* [12] contain both homodimeric and tetrameric hemoglobins. Besides *B. virescens*, a heterodimeric hemoglobin is occurred only in the clam *Noetia ponderosa* [13] and might represent a prototype of tetrameric hemoglobin. The two constituent chains, I and II, of *B. virescens* hemoglobin have rather different

FIG. 6. HPLC purification of CNBr fragments of *B. lima* two-domain hemoglobin. The column (Cosmosil $5C_{18}$-300) was equilibrated with solvent A (2% acetonitrile in 0.1% TFA) and eluted with a linear gradient of 0 to 100% solvent B (80% acetonitrile in 0.1% TFA) over 60 min. Flow rate, 1 ml/min. A, fraction 1 in Fig. 5; B, fraction 2; C, fraction 3.

Table 2. Amino acid compositions of CNBr, trypsin and lysyl endopeptidase

	CN1	CN1T1	CN1T1L1	CN1T1L2	CN1T2	CN1T3	CN2
Asp	4.0(4)	2.0(2)	1.0(1)	1.1(1)	1.0(1)	1.0(1)	2.0(2)
Thr	1.7(2)	0.9(1)		1.0(1)		0.8(1)	
Ser	1.6(2)	0.9(1)	0.9(1)			0.8(1)	
Glu	4.1(4)	3.9(4)	1.0(1)	3.1(3)			1.1(1)
Gly							2.7(3)
Ala	1.1(1)	1.0(1)		1.1(1)			0.9(1)
Cys							
Val	3.0(3)	1.9(2)	1.0(1)	1.1(1)		1.0(1)	1.0(1)
Met							
Ile	1.8(2)	0.9(1)		1.0(1)	0.9(1)		
Leu	1.1(1)				1.0(1)		1.2(1)
Tyr							
Phe							
Lys	1.9(2)	1.9(2)	1.0(1)	1.1(1)			
His							
Arg	1.0(1)				0.9(1)		1.1(1)
Pro	1.0(1)	1.0(1)		1.0(1)			
Trp	+ (1)					+ (1)	
Hse**	+ (1)					+ (1)	+ (1)
Total	25	15	5	10	4	6	11
Position	D1$	D1	D1	D1	D1	D1	D1
	1–25	1–15	1–5	6–15	16–19	20–25	26–36
Yield(%)	37.6	37.2	37.0	37.0	37.2	27.7	12.1

*Due to heterogeniety. **Homoserine. $Domain 1. #Domain 2.

amino acid compositions; the SΔQ value [14] of chain I is 43.3 and that of chain II is 83.4, suggesting that they have rather different amino acid sequences. N-Terminal 20 residues of chain I shows homology (35% identity) with those of homodimeric chains from *Anadara* [11] and *Scapharca* [12], but shows little homology with the polymeric chain from *B. lima* (see Table 3).

B. virescens hemoglobin tended to polymerize. For example, when the aged sample was applied to the gel filtration column under the same conditions shown in Figure 1, several aggregates with higher *Mr* were newly emerged (data not shown). But HPLC analyses of the aggregates in the presence of a reducing agent gave the same elution profile as shown in Figure 4A. Since there is no cysteine residue in chain I, that in chain II must be responsible for the polymerization of *B. virescens* hemoglobin.

B. virescens chain I seems to contain only two histidine residues, which most likely correspond to the proximal and distal histidines (Table 1). Compared with vertebrate hemoglobins, molluscan globins have a low histidine content; *Anadara* dimeric and tetrameric hemoglobins have 2-3 histidines per chain [11]; *Scapharca* dimeric hemoglobin, 2 histidines [12]; *Aplysia* myoglobin, one histidine [15]; and *Dolabella* myoglobin, one histidine [8].

On the other hand, *B. lima* contained tetrameric and polymeric hemoglobins, like *B. reeveana* [4]. The tetrameric hemoglobin was composed of two myoglobin-like chains (A and B), while the polymeric hemoglobin was composed of unusual two-domain polypeptide chain (P). The amino acid compositions of chains A, B and P were rather different, the SΔQ values being 50.9, 131.3 and 98.9, respectively (Table 1). Chain B was especially rich in alanine and had lower content of glutamic

peptides of *Barbatia lima* two-domain hemoglobin

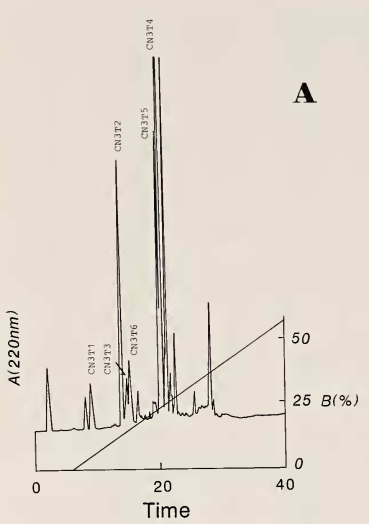
CN3	CN3T1	CN3T2	CN3T3	CN3T4	CN3T5	CN3T6	L1	L2
4.1(4)		1.1(1)		1.9(2)	2.0(2)	1.0(1)	1.1(1)	1.2(1)
2.2(2.5)*	0.9(1)	0.7(0.5)*		0.9(1)	0.9(1)		1.6(2)	0.9(1)
5.6(6)	1.0(1)	2.8(3)		1.1(1)	1.2(1)	0.9(1)	4.1(4)	3.9(4)
4.0(4)	1.0(1)		1.0(1)	1.2(1)	1.2(1)	1.0(1)	1.1(1)	1.2(1)
2.4(2.5)*		1.2(1.5)*		1.0(1)	1.0(1)		1.2(1)	1.9(2)
3.6(4)	1.0(1)	1.0(1)		0.8(1)	0.9(1)	1.0(1)	2.0(2)	2.0(2)
							0.8(1)	0.9(1)
2.5(3)		1.0(1)	0.9(1)	0.8(1)	0.8(1)		0.9(1)	0.9(1)
2.3(2)			1.0(1)			1.0(1)		
2.5(2.5)*		1.0(1)		1.1(1)	1.9(2)		1.1(1)	1.2(1)
2.5(2.5)*	1.0(1)		1.0(1)	0.8(1)			1.0(1)	1.0(1)
0.9(1)		0.9(1)					1.1(1)	1.0(1)
+ (1)				+ (1)	+ (1)			
+ (1)						+ (1)		
36	5	10	4	11	11	6	16	16
D2#	D2	D2	D2	D2	D2	D2	D2	D2
1–36	1–5	6–15	16–19	20–30	20–30	31–36	0–15	0–15
38.9	22.5	26.0	22.5	11.5	4.6	6.9	27.7	11.9

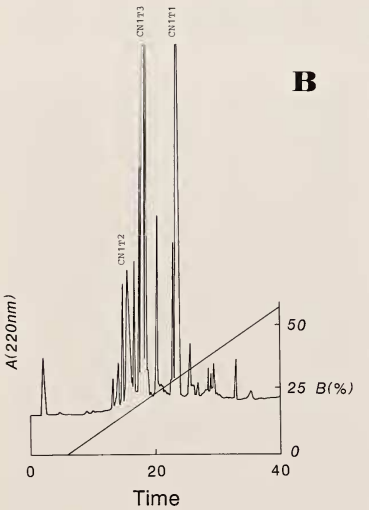
acid. The composition of chain A resembled that of *B. virescens* chain I, except glycine residue. Calculated from the *Mr* 32,000 of chain P, the chain has about 290 amino acids, just two times as many as typical globin.

The physiological role of *B. lima* two-domain hemoglobin is uncertain. Although *Barbatia* two-domain chain surely has two oxygen-binding sites [5] and can bind oxygen [4], the oxyhemoglobin isolated undergoes very rapid autoxidation and tends to precipitate. Although the two-domain hemoglobin is also found in the hemolymph of the water fleas *Daphnia* and *Moina* [16], unlike *Barbatia* two-domain hemoglobin, *Daphnia* hemoglobin is very resistant to autoxidation (Suzuki and Kobayashi, unpublished result). Therefore *Barbatia* two-domain hemoglobin seems to be rather disadvantageous as oxygen carrier protein.

To examine the evolutionary origin of two-domain hemoglobin, we determined the N-terminal sequence of 36 residues of each domain of *B. lima* chain P. The partial sequences are aligned with that of *B. reeveana* domain 1 [5] in Figure 9. Of the three sequences, 24 out of 36 residues (67%) appear to be invariant, the highest homology being found between *B. lima* domain 1 and *B. reeveana* domain 1 (83% identity). The sequence homology between *B. lima* two domains was also high (75% identity). The presence of N-terminal Met of *B. reeveana* domain I is not sure, due to DNA sequencing. N-Terminus of *B. lima* domain 1 was blocked by an acetyl group, and the second domain appeared to start at methionine residue. From these results, it seems that *Barbatia* two-domain chain resulted from gene duplication and the following loss of a stop codon of the first domain.

We have shown that the two closely related





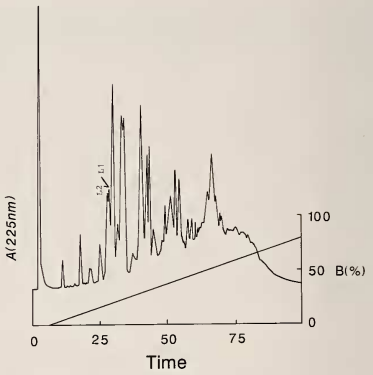


FIG. 8. HPLC separation of lysyl endopeptidase peptides of *B. lima* two-domain hemoglobin. Conditions are as for Fig. 6.

clams *B. virescens* and *B. lima* have quite different types of hemoglobins. It is usual that the closely related species have a similar hemoglobin in sequence and also in quaternary structure. For example, the related clams *Anadara trapezia, Anadara broughtonii* and *Scapharca inaequivalvis* have both dimeric and tetrameric hemoglobins, the sequence homology between corresponding chains being very high (over 85%) [11]; the related sea hares *Aplysia* and *Dolabella* have a monomeric myoglobin, the sequence homology being over 70% [8]. However, the SΔQ values (Table 1) of *B. virescens* and *B. lima* chains do not suggest strong sequence homology between any chains. At present, we have no idea to explain this unusual phenomenon. But it is likely that this results from a taxonomical confusion.

N-Terminal amino acid sequences of invertebrate hemoglobins are compared in Table 3, which includes those of the phylum Echinodermata (sea cucumbers *Molpadia* [17] and *Paracaudina* (Suzuki, unpublished)), of Annelida (sea worms *Glycera*

FIG. 7. HPLC separation of tryptic peptides of *B. lima* CN3 (A) and CN1 (B). Conditions are as for Fig. 6.

TABLE 3. N-Terminal sequences of invertebrate hemoglobins

		−10 −5 0 5 10 15 pre A — A
Vertebrata	Human tetramer beta (18)	V H L T P E E K S A V T A L W G K
	Lamprey monomer (18)	P I V D S G S V A P L S A A E K T K I R S A W A P
Echinodermata	*Molpadia* D (17)	Ac–G Q(T S A)F Q S V G D L T L A E K D L I R S T W D N
	Paracaudina I (unpublished)	Ac–G G T L A I Q S H G D L T L A Q K K I V R K T W H Q
Annelida	*Glycera* monomer (18)	G L S A A Q R Q V I A A T W K D
	Tylorrhynchus IIA (9)	S S D H C G P L Q R L K V K Q Q W A K
	Lumbricus I (19)	E C L V T E G L K V K L Q W A S
	Pheretima I (unpublished)	D C N T L K R F K V K H Q W Q Q
Vestimentifera	*Lamellibrachia* II-I (unpublished)	D C N I L Q R L K V K M Q W A K
Arthropoda	*Chironomus* I (18)	G P S G D Q I A A A K A S W N T
	Artemia domain E1 (20)	E R V D P I T G L S G L E K N A I L D T W G K
	Moina domain 2 (unpublished)	A P E D L V D P E T R L S G I H K
Mollusca	*Anadara* tetramer beta (11)	Ac–S T V A E L A N A V V S N A D Q K D L L R L S W G V
	dimer (11)	P S V Q D A A A Q L T A D V K K D L R D S W K V
	Barbatia virescens dimer I (This work)	P S V — — A A A V S A V T N K D V A Q E I W
	Barbatia reeveana domain 1 (5)	(M)S V S A K L D E V T Q P A N K N L I R S T W N M
	Barbatia lima domain 1 (This work)	Ac–S V E D K I E E V T Q P A N K N L I R S T W N V
	domain 2 (This work)	M G V T E R I E E V T Q P A N K G L I R E T W N I
	Calyptogena chain I (unpublished)	V S A N D I K N V Q — D T W G K
	chain II (unpublished)	V S Q A D I A A V Q — T S W R R

Ac, an acetyl group. Helical segments by Chou and Fasman method [21] are marked by ····.

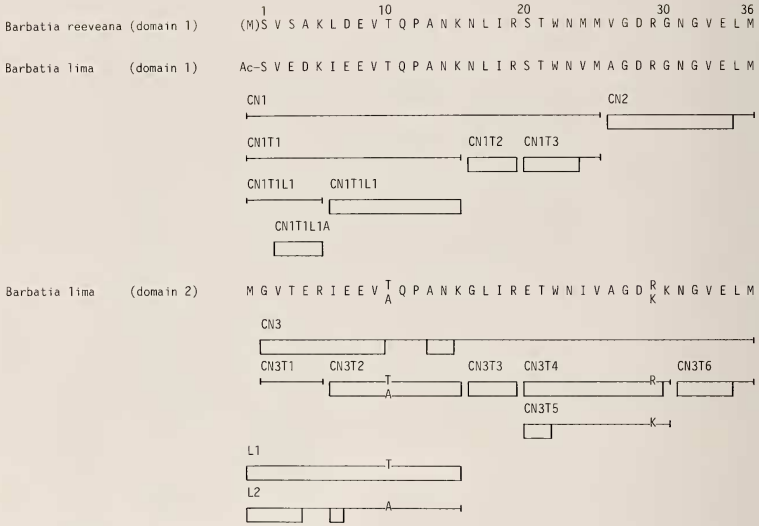


FIG. 9. Summary of data to establish the N-terminal sequences of 37 residues of each domain of *B. lima* two-domain hemoglobin. Manual Edman degradation (□) was employed for sequence determination. The domain 1 sequence of *B. reeveana* hemoglobin [5] is aligned together. CN, a CNBr peptide; T, a tryptic peptide; L, a lysyl endopeptidase peptide; A, a acylamino-acid-releasing enzyme peptide.

[18] and *Tylorrhynchus* [9], and earthworms *Lumbricus* [19] and *Pheretima* (Suzuki, unpublished)), of Vestimentifera (the deep-sea cold seep tube worm *Lamellibrachia* sp. (undescribed)(Suzuki, Takagi and Ohta, unpublished), of Arthropoda (*Chironomus* [18], *Artemia* [20] and *Moina* (Suzuki and Kobayashi, unpublished)) and of Mollusca (*Anadara* [11], *Barbatia* (this work, [5]) and the deep-sea cold seep clam *Calyptogena soyoae* (Suzuki, Takagi and Ohta, unpublished)). These sequences were aligned by two conservative residues, Lys at position 7(A5) and Trp at 14 (A12). As shown in Table 3, many invertebrate sequences have the N-terminal extension (pre-A segment) when compared with most of vertebrate globins. Since the extension of *Scapharca* hemoglobin has been shown to form an additional helix (pre-A helix), we examined the secondary structure of N-terminal sequences of invertebrate hemoglobins by Chou-Fasman method [21]. Consequently, A-helix was predicted to most of all globins, and the pre A-helix was also predicted to the globins from *Molpadia, Moina, Anadara* and *Barbatia.* Anyway, the pre-A segment, which is distributed widely in invertebrate hemoglobin sequences, might represent a prototype of globin sequence.

ACKNOWLEDGMENTS

We are indebted to Mr Zenji Imoto for the supply of *B. lima.*

REFERENCES

1 Grinich, N. P. and Terwilliger, R. C. (1980) The quaternary structure of an unusually high mol. wt intracellular hemoglobin from the bivalve mollusc

Barbatia reeveana. Biochem. J., **189**, 1–8.

2 Terwilliger, R. C. and Terwilliger, N. B. (1985) Molluscan hemoglobins. Comp. Biochem. Physiol., **81B**, 255-261.

3 Royer Jr, W. E., Love, W. E. and Fenderson, F. F. Cooperative dimeric and tetrameric clam haemoglobins are novel assemblages of myoglobin folds. (1985) Nature, **316**, 277–280.

4 Grinich, N. P. Terwilliger, R. C. and Terwilliger, N. B. (1986) Oxygen equilibria and structural characteristics of the tetrameric and polymeric intracellular hemoglobins from the bivalve mollusc *Barbatia reeveana.* J. Comp. Physiol., B 156, 675–682.

5 Riggs, A. F., Riggs, C. K., Lin, R.-J, and Domdey, H. (1986) Cloning of the cDNA for the globin from the clam, *Barbatia reeveana.* In "Invertebrate Oxygen Carriers". Ed. by B. Linzen, Springer-Varlag, Berlin/Heidelberg, pp. 473–476.

6 Shelton, J. B., Shelton, J. R., Schroeder, W. A. and DeSimone, J. (1982) Detection of Hb-Papio B, a slight mutation of the baboon beta chain, by high performance liquid chromatography. Hemoglobin, **6**, 451–464.

7 Ashauer, H., Weber, R. E. and Braunitzer, G. (1985) The primary structure of the hemoglobin of the dog fish shark (*Squalus acanthias*). Biol. Chem. Hoppe-Seyler, **366**, 589–599.

8 Suzuki, T. (1986) Amino acid sequence of myoglobin from the mollusc *Dolabella auricularia.* J. Biol. Chem., **261**, 3692–3699.

9 Suzuki, T. and Gotoh, T. (1986) The complete amino acid sequence of giant multisubunit hemoglobin from the polychaete *Tylorrhynchus heterochaetus.* J. Biol. Chem., **261**, 9257–9267.

10 Furuta, H., Ohe, M. and Kajita, A. (1977) Subunit structures of hemoglobins from erythrocytes of the blood clam, *Anadara broughtonii.* J. Biochem., **82**, 1723–1730.

11 Gilbert, A. T. and Thompson, E. O. P. (1985) Amino acid sequence of the beta-chain of the tetrameric hemoglobin of the bivalve mollusc, *Anadara trapezia.* Aust. J. Biol. Sci., **38**, 221–236.

12 Petruzzelli, R., Goffredo, B. M., Barra, D., Bossa, F., Boffi, A., Verzili, D., Ascoli, F. and Chiancone, E. (1985) Amino acid sequence of the cooperative homodimeric hemoglobin from the mollusc *Scapharca inaequivalvis* and topology of the intersubunit contacts. FEBS Lett., **184**, 328–332.

13 San George, R. C., and Nagel, R. L. (1985) Dimeric hemoglobins from the acid blood clam, *Noetia ponderosa.* J. Biol. Chem., **260**, 4331–4337.

14 Cornish-Bowden, A. (1983) The amino acid compositions of proteins are correlated with their molecular sizes. Biochem. J., **213**, 271–274.

15 Suzuki, T., Takagi, T., and Shikama, K. (1981) Amino acid sequence of myoglobin from the mollusc *Aplysia kurodai.* Biochim. Biophys. Acta, **669**, 79–83.

16 Kobayashi, M. and Hoshi, T. (1984) Analysis of respiratory role of haemoglobin in *Daphnia magna.* Zool. Sci., **1**, 523–532.

17 Mauri, F. C. (1985) Ph. D. Dissertation, Unviersity of Texas, Austin, Texas

18 Kleinschmidt, T. and Sgouros, J. G. (1987) Hemoglobin sequences. Biol. Chem. Hoppe-Seyler, **368**, 579–615.

19 Shishikura, F., Snow, J. W., Gotoh, T., Vinogradov, S. N. and Walz, D. A. (1987) Amino acid sequence of the monomer subunit of the extracellular hemoglobin of *Lumbricus terrestris.* J. Biol. Chem., **262**, 3123–3131.

20 Moens, L., Van Hauwaert, M. L., Geelen, D., Verpooten, G. and Van Beeumen, J. (1986) The amino acid sequence of a structural unit isolated from the high molecular weight globin chains of *Artemia* sp. In "Invertebrate Oxygen Carrier". Ed. by B. Linzen, Springer-Verlag/Berlin/Heidelberg, pp. 81–84.

21 Chou, P. Y. and Fasman, G. D. (1978) Empirical predictions of protein conformation. Ann. Rev. Biochem., **47**, 251–276.